Journal of the Nigerian Society of Physical Sciences

Modified Gradient Flow Method for Solving One-Dimensional Optimal Control Problem Governed by Linear Equality Constraint

Olusegun Olotu[a], Charles Aladesaye[b], Kazeem Adebowale Dawodu[a,*]

[a]*Department of Mathematical Sciences,The Federal University of Technology Akure, Nigeria*
[b]*Dept. of Mathematics, School of Science, College of Education, Ikere-Ekiti, Ekiti State, Nigeria*

Abstract

This study presents a computational technique developed for solving linearly constraint optimal control problems using the Gradient Flow Method. This proposed method, called the Modified Gradient Flow Method (MGFM), is based on the continuous gradient flow reformulation of constrained optimization problem with three-level implicit time discretization scheme. The three-level splitting parameters for the discretization of the gradient flow equations are such that the sum of the parameters equal to one ($\theta_1 + \theta_2 + \theta_3 = 1$). The Linear and quadratic convergence of the scheme were analyzed and were shown to have first order scheme when each parameter exist in the domain [0, 1] and second order when the third parameter equal to one. Numerical experiments were carried out and the results showed that the approach is very effective for handling this class of constrained optimal control problems. It also compared favorably with the analytical solutions and performed better than the existing schemes in terms of convergence and accuracy.

1. Introduction

Many authors, in literature, have attempted the class of linearly constrained optimal control problems using the classical analytical techniques of the Hamilton-Jacobi-Bellman (HJB) or "Pontryagin Maximum Principle" by Pontryagin *et al.* in [14]; though this approach may proof difficult in some cases when inclusive of non-differentiable delay terms. The numerical approach has also been fully handled by several authors like Betts in [4], Olotu and Dawodu in [11, 12], Adekunle [1] and Akeremale [2], using the direct numerical methods. This method deploys the "first discretized - then - optimize" technique which is known for its amenability to well-structured discretization schemes and optimization solvers after the unconstrained formulation of the control problem using any of the penalty function methods or method of multipliers. The direct method of nonlinear programming formulation tends to eradicate the problem of ill-conditioning associated with control problems with delay dynamical systems.

However, this research paper seeks for a better approach in the

---

*Corresponding author tel. no: +2348032541380
Email address: dawodukazeem@futa.edu.ng (Kazeem Adebowale Dawodu)

formulation of the control operators in the optimization of the discretized optimal control problem. The gradient flow method was first applied to nonlinear programming problems by Evtushenko [5] and later improved by Evtushenko and Zhadan [6]; while the application to unconstrained problems was by Behrman [3] and the unified approach to nonlinear optimization problem was by Wang *et al.* [15]. The unified gradient flow algorithm exhibits linear convergence for $\theta \in [0, 1)$ and quadratic convergence at $\theta = 1$. As an improvement over [15], a modified gradient flow method MGFM was proposed using the gradient formula for Lagrange multiplier formulation. The development of this method requires a continuous gradient flow approach based on a three level implicit time discretization scheme with splitting parameter θ_1, θ_2 and θ_3. The convergence of the solution of the modified discretized gradient flow equation to a local minimum of the original problem, either linearly or quadratically depending on the choice of θ, will as well be demonstrated. However, the technique can also be adapted to handle inequality constraint problems by introducing slack variables. Three test problems were solved and their numerical results presented in comparison with existing methods. This is to demonstrate the effectiveness of the method in solving constrained nonlinear programming problems with either continuous variables or a mixture of continuous and discrete variables.

The structure of this paper is as follows: Section 3 is dedicated to the idea of gradient flow approach for unconstrained optimization while Section 4 presents the modified gradient flow algorithm for constrained optimization in Section 2 using the 3-level splitting parameters and as well as the convergence analysis. The Algorithm was experimented on two numerical examples and comparison made with the Conjugate Gradient Method (CGM) in Section 4.

2. Statement of problem

The general form of the One-Dimensional Optimal Control Problem is given as

$$\text{Minimize } J(x, u) = \int_{t_0}^{T} F(t, x(t), u(t))dt, \tag{1}$$

$$\text{Subject to } \dot{x}(t) = f(t, x(t), u(t)), \ t \in [t_0, T], \tag{2}$$

$$x(t_0) = x_0, \tag{3}$$

where $x(t) \in C^1[t_0, T] \in \mathbb{R}$, $u(t) \in L^2[t_0, T] \in \mathbb{R}$, $F : [t_0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ and $f : [t_0, T] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$.

Thus, $x(t)$ is the state variable which describes the state of the system at any point in time and $u(t)$ is the control variable in the optimization problem. However, for linear-quadratic problem, the objective functional is written as $F(t, x(t), u(t)) = (px^2(t) + qu^2(t))$ while the linear differential constraint is written as $f(t, x(t), u(t)) = ax(t) + bu(t)$, where $p, q, a, b \in \mathbb{R}$ and $p, q > 0$.

3. Research Methodology

This paper considered both the theoretical and numerical analyses of the problem in the development of the algorithm for solving eqns. (1) to (3). The discrete form of the quadratic objective functional and the linear differential constraint are constructed using the Trapezoidal and Euler method respectively. The recurrence relation for both the objective functional and constraint were developed to help construct their respective matrix operators. Furthermore, the symmetry and positive definite properties of the matrix operators were analyzed to guarantee their consistency and invertibilty in the developed MGFM.

3.1. Preamble

Let $f : X \to \mathbb{R}$ be a functional such that the solution $x(t)$ is a minimizer of f. The gradient flow method, with an initial point $x(0) \in X$, seeks to find a minimizer of f by the curve $x(t)$ defined by the differential equation

$$\frac{dx(t)}{dt} = -\nabla f(x(t)), \quad x(0) = x_0, \tag{4}$$

where ∇f is called the gradient of f and the solution is the called the integral curve such that at each instance proceeds in the direction of the steepest descent of f. In an unconstrained optimization problem, the equilibrium points of the gradient flow are the zeros of ∇f, which are exactly the minimizers of f. The gradient flow solves the problem min $f(x)$, $x(0) = x_0$ such that in every trajectory $x(t)$ of the gradient flow, then $f(x(t)) \to p^*$ for p^* is the minimizer of f.

For clarity, the derivative $x'(t)$ is descretized using the forward Euler step which yields

$$x^{k+1} := x^k - h\nabla f(x^k) \tag{5}$$

or the backward discretization scheme which yields

$$x^{k+1} := x^k - h\nabla f(x^{k+1}). \tag{6}$$

The backward Euler method can be applied for the numerical integration of gradient flow differential equation, also known as the *Proximal minimization method* expressed thus:

$$x^{k+1} := \mathbf{Prox}_{h,f}(x^k).$$

However, the proximal gradient flow algorithm can be interpreted as gradient flows,

$$\frac{dx(t)}{dt} = -\nabla f(x(t)) - \nabla g(x(t)). \tag{7}$$

where g is differentiable. Using the forward-backward integration of the gradient flow, eqn. (7) can be expressed as

$$x^{k+1} = x^k - h\nabla f(x^k) - h\nabla g(x^{k+1}),$$

$$x^{k+1} := (I + h\nabla g)^{-1}(I - h\nabla f)x^k,$$

where the *forward-backward splitting method* above numerically integrates the gradient flow differential equations.

The time-stepping discretization of eqns. (5) and (6) is expressed as follows:

$$x^{k+1} := x^k - h\left[(1-\theta)\nabla f(x^k) + \theta\nabla f(x^{k+1})\right], \qquad (8)$$

where $\theta \in [0,1]$ is a parameter. When $\theta = 0$, the above discretization is the explicit Euler's scheme and on the other hand, the implicit backward Euler when $\theta = 1$.

3.2. Discretization of the Objective function

The discretization of the objective function in eqn. (1) above, the Trapezoidal rule will be used given as;

$$\int_{t_0}^{T} f(x)dx = \frac{h}{2}\sum_{i=0}^{N-1}[f(x_i) + f(x_{i+1})], \qquad (9)$$

Expanding eqn. (9) above, yields;

$$J(x,u) \approx \frac{h}{2}p\sum_{i=0}^{N-1}(x_i^2 + x_{i+1}^2) + \frac{h}{2}q\sum_{i=0}^{N-1}(u_i^2 + u_{i+1}^2), \qquad (10)$$

where for $t_i = t_0 + ih$ then $x(t_i) \approx x(t_0 + ih) = x_i$, $u(t_i) \approx u(t_0 + ih) = u_i$ for $i = 1, 2, ..., N$ and $N = (T - t_0)/h$. Rearranging eqn. (10) into matrix form yields;

$$\text{Minimize} \quad J(Z) = \frac{1}{2}Z^T M Z + C, \qquad (11)$$

where $Z = (x_1, x_2, ...x_N, u_0, u_1, ..., u_N)^T \in \mathbb{R}^{2N+1}$, $C = \frac{1}{2}x_0^2 hp \in \mathbb{R}$ and $M \in \mathbb{R}^{(2N+1)\times(2N+1)}$ is a $(2N+1)\times(2N+1)$ matrix operator with diagonal entries m_{ij} as expressed below:

$$[m_{ij}] = \begin{cases} 2hp, & 1 \le i \le N-1, & j = i, \\ hp, & i = N, & j = i, \\ hq, & i = N+1, 2N+1, & j = i, \\ 2hq, & N+2 \le i \le 2N, & j = i, \\ 0, & & \text{elsewhere} \end{cases} \qquad (12)$$

3.3. Discretization of the Constraint function

The differential constraint $\dot{x}(t) = f(t, x(t), u(t)) = ax(t) + bu(t)$ was discretized by slotting it into the Euler's scheme $x_{i+1} = x_i + \dot{x}h$ with $t_i = t_0 + ih$, to arrive at

$$x_{i+1} = cx_i + du_i, \qquad (13)$$

where $c = 1 + ha$, $d = hb$ and $i = 0, 1, 2, ..., N - 1$. And upon expansion and re-arrangement of eqn. (13) into matrix form yields $GZ = k$ express as

$$g(Z) = GZ - k = 0, \qquad (14)$$

where Z is of dimension $(2N + 1) \times 1$ and G is of dimension $N \times (2N + 1)$ as shown below;

$$G = \begin{bmatrix} 1 & 0 & \cdots & 0 & -d & 0 & \cdots & \cdots & 0 \\ -c & 1 & \ddots & \vdots & 0 & -d & 0 & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 & \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & -c & 1 & 0 & \cdots & 0 & -d & 0 \end{bmatrix}$$

and

$$k = \begin{bmatrix} cx_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

with the entries defined as follows: $g_{ij} = 1$ for $1 \le i \le N$ and $j = i$; $g_{ij} = -d$ for $1 \le i \le N$ and $j = N + i$; $g_{ij} = -c$ for $2 \le i \le N$ and $j = i - 1$ while $g_{ij} = 0$ otherwise. Hence, eqns. (11) and (14) give the quadratic representation for the constrained minimization problem as

$$\text{Minimize} \quad J(Z) = \frac{1}{2}Z^T M Z + C,$$
$$\text{subject to} \quad GZ - k = 0. \qquad (15)$$

3.4. Lagrangian formulation of the Constraint problem

The unconstrained form of the optimal control problem can be obtained using the Lagrangian function on eqn. (15) as stated below:

$$L(Z, \lambda) = \frac{1}{2}Z^T M Z + C + \lambda^T(GZ - k), \qquad (16)$$
$$L(Z, \lambda) = J(Z) + \lambda^T g(Z), \qquad (17)$$

where $\lambda = (\lambda_1, \lambda_2, ..., \lambda_N)^T \in \mathbb{R}^N$ is the Lagrangian multiplier. The pair of points (Z^*, λ^*) is said to be a stationary pair points of eqn. (17) if the following first order necessary optimality conditions hold:

$$L_Z(Z^*, \lambda^*) = J_Z(Z^*) + g_Z^T(Z^*)\lambda^* = 0, \qquad (18)$$
$$L_\lambda(Z^*, \lambda^*) = g(Z^*) = 0, \qquad (19)$$

where

$$L_Z(Z, \lambda) = \left(\frac{\partial L(Z, \lambda)}{\partial x_1}, \cdots, \frac{\partial L(Z, \lambda)}{\partial u_N}\right)^T,$$
$$L_\lambda(Z, \lambda) = \left(\frac{\partial L(Z, \lambda)}{\partial \lambda_1}, \cdots, \frac{\partial L(Z, \lambda)}{\partial \lambda_N}\right)^T,$$
$$J_Z(Z) = \left(\frac{\partial J(Z)}{\partial x_1}, \cdots, \frac{\partial J(Z)}{\partial u_N}\right)^T,$$

denote the *gradient vectors* of L and J with the dimensions $(2N + 1) \times 1$, $N \times 1$ and $(2N + 1) \times 1$ respectively. Similarly, we

have the Jacobian of g with dimension $N \times (2N+1)$ represented below:

$$g_Z(Z) = \begin{bmatrix} \frac{\partial g_1(Z)}{\partial x_1} & \frac{\partial g_1(Z)}{\partial x_2} & \cdots & \frac{\partial g_1(Z)}{\partial u_N} \\ \frac{\partial g_2(Z)}{\partial x_1} & \frac{\partial g_2(Z)}{\partial x_2} & \cdots & \frac{\partial g_2(Z)}{\partial u_N} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial g_N(Z)}{\partial x_1} & \frac{\partial g_N(Z)}{\partial x_2} & \cdots & \frac{\partial g_N(Z)}{\partial u_N} \end{bmatrix}. \quad (20)$$

In order to fulfill the above optimality conditions, the continuous gradient flow reformulation with the given initial condition Z_0 is derived as follows;

$$\frac{dZ}{dt} = -L_Z(Z(t), \lambda(Z(t))), \; Z(0) = Z_0, \quad (21)$$

$$\frac{dZ}{dt} = \varphi(Z). \quad (22)$$

Substituting eqn. (18) into (21) yields,

$$\frac{dZ}{dt} = -[J_Z(Z(t)) + g_Z^T(Z(t))\lambda^0(Z(t))], \; Z(0) = Z_0 \quad (23)$$

such that the choice of $\lambda^0(Z)$ is to satisfy

$$L_Z(Z, \lambda^0(Z)) = \tau g_Z^T(Z)g(Z), \quad (24)$$

in the least square sense, where $\tau > 0$ is a parameter.
Substituting eqn. (18) into (24) yields,

$$J_Z(Z) + \lambda^0(Z)g_Z^T(Z) = \tau g_Z^T(Z)g(Z). \quad (25)$$

Obtaining the least squares solution of $\lambda^0(Z)$ in en. (25) yields,

$$\lambda^0(Z)g_Z^T(Z) = \tau g_Z^T(Z)g(Z) - J_Z(Z), \quad (26)$$

$$\lambda^0(Z) = [g_Z^T(Z)]^\dagger \{\tau g_Z^T(Z)g(Z) - J_Z(Z)\}, \quad (27)$$

where the superscript '†' refers to the pseudo inverse of $g_Z^T(Z)$ in Layton [9].

Theorem 3.1. *The constraint equations for the discretized optimal control problem $g(Z)$ satisfies the linearly independent constraint qualification (LICQ) for any feasible point Z^*. The LICQ states that the gradient of the active inequality constraints and the gradient of the equality constraints are linearly independent at Z^*. eyiu*

LICQ holds if $g_{1Z}(Z), g_{2Z}(Z), ..., g_{NZ}(Z)$ are linearly independent for any feasible point Z^*. Thus

$$\alpha_1 g_{1Z}(Z) + \alpha_2 g_{2Z}(Z) + ... + \alpha_N g_{NZ}(Z) = 0 \quad (28)$$

implies that $\alpha_i = 0$. Hence, the dimension of the row space of G must be N.
Applying the row-reduction process to G in eqn. (3.3), we have $\bar{G}_* = [I|V]$ written as
$\bar{G}_* =$

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & -d & 0 & \cdots & \cdots & 0 \\ 0 & 1 & \ddots & \vdots & -cd & -d & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 & -c^{N-1}d & \cdots & -cd & -d & 0 \end{bmatrix}, (29)$$

where $I \in \mathbb{R}^{N \times N}$ and $V \in \mathbb{R}^{N \times (N+1)}$. The entries g_{ij}^* of $\bar{G}_*$ can be defined as follows:

$$[g_{ij}^*] = \begin{cases} 1, & 1 \le i \le N, \; j = i; \\ -c^k d, & k+1 \le i \le N, \; j = N+i-2; \\ & \text{for } k = 1, 2, ..., N-1, \\ 0, & elsewhere. \end{cases} \quad (30)$$

Thus, rank of $\bar{G}_*$ equals rank of G and $\alpha_i = 0$ are the parameters that make $g_{1Z}(Z), g_{2Z}(Z), ..., g_{NZ}(Z)$ linearly independent for any feasible point Z^* and $G(Z)$ satisfies LICQ. From theorem (3.1), we have for any feasible point Z^* that $\bar{G}_* \in \mathbb{R}^{N \times (2N+1)}$ is re-structured into a $N \times N$ non-singular (invertible) Gram matrix G_* given as

$$G_* = g_Z(Z)g_Z^T(Z). \quad (31)$$

Multiplying both sides of eqn. (25) by $g_Z(Z)$ arrives at

$$g_Z(Z)(J_Z(Z) + \lambda^0(Z)g_Z^T(Z)) = \tau g_Z(Z)g_Z^T(Z)g(Z). \quad (32)$$

Therefore, with $G_*(Z) = g_Z(Z)g_Z^T(Z)$, eqn. (32) becomes

$$g_Z(Z)J_Z(Z) + g_Z(Z)\lambda^0(Z)g_Z^T(Z) = \tau G_* g(Z)$$

and simplified to get

$$G_* \lambda^0(Z) = \tau G_* g(Z) - g_Z(Z)J_Z(Z). \quad (33)$$

Pre-multiplying eqn. (33) by G_*^{-1} yields

$$G_*^{-1} G_* \lambda^0(Z) = G_*^{-1} \tau G_* g(Z) - G_*^{-1} g_Z(Z)J_Z(Z), \quad (34)$$

and solving for $\lambda^0(Z)$ to obtain

$$\lambda^0(Z) = \tau g(Z) - G_*^{-1} g_Z(Z)J_Z(Z), \quad (35)$$

where eqns. (27) and (35) are equivalent. Consequently, eqn. (27) can be express as

$$\begin{aligned} \lambda^0(Z) &= (g_Z^T(Z))^\dagger [\tau g_Z^T(Z)g_Z(Z) - J_Z(Z)], \\ &= \tau g_Z(Z) - G_*^{-1} J_Z(Z)g_Z(Z), \end{aligned}$$

where $G_*^{-1} = [g_Z^T(Z)g_Z(Z)]^{-1}$, hence the pseudo-inverse of $[g_Z^T(Z)]$ is $[g_Z^T(Z)]^\dagger = G_*^{-1} g_Z(Z)$.
However, eqn. (35) is true when Z is sufficiently close to an equilibrium point Z^*. The choice of $\lambda^0(Z)$ has the advantage of reducing the complexity in the evaluation of the gradient of λ^0 with respect to Z. Thus substituting $\lambda^0(Z)$ in eqn. (35) above into eqn. (23) yields;

$$\begin{aligned} \dot{Z} &= \varphi(Z) = -(J_Z(Z) + g_Z^T(Z)[\tau g(Z) \\ &\quad - G^{-1}(Z)g_Z(Z)J_Z(Z)]). \end{aligned} \quad (36)$$

In testing the convergence of the gradient flow formulation expressed in eqns. (23) and (24), we then considered the LICQ theorem below.

Theorem 3.2. *Suppose that (Z^*, λ^*) is a stationary point, then*

1. *LICQ holds at Z^*.*
2. *For any $\bar{Z}$ satisfying $g_Z(Z^*)\bar{Z} = 0$, then,*
   $\bar{Z}^T L_{ZZ}(Z^*, \lambda^*)\bar{Z} > 0$ where $\bar{Z} = Z - Z^*$.

It indicates that the pair $(Z(t), \lambda^0(Z(t)))$ tends to (Z^*, λ^*) as $t \to \infty$ provided the starting point Z_0 is sufficiently close to point Z^*.

Proof:

We start by deploying the ***Poincare-Lyapunov*** stability theorem stated in Haddad *et al.* [8] and Morris [10] thus:

Theorem 3.3 (Liapunov Stability). *Let X^* be an equilibrium point for $X' = J(X)$. Let $L : \Omega \to R$ be a differentiable function defined on an open set Ω containing X^*. Suppose further that*

1. *$L(X^*) = \Omega$ and $L(X) > 0$ if $X \neq X^*$;*
2. *$\dot{L} \leq 0$ in $\Omega - X^*$; then X^* is stable. Furthermore, if L also satisfies*
3. *$\dot{L} < 0$ in $\Omega - X^*$, then X^* is asymptotically stable,*

It then requires we show that Z^* is an asymptotically stable point for $\dot{Z} = \varphi(Z)$ if $\varphi(Z)$ is continuously differentiable and the original of the linearized system

$$\dot{y} = \varphi_Z(Z^*)y, \tag{37}$$

where $y := Z - Z^*$, is exponentially stable (i.e. all eigenvalues of $\varphi_Z(Z^*)$ are strictly negative). It was however noted that at $Z^*,(Z^*, \lambda^0(Z^*))$ satisfies eqns. (18) and (19) because the right hand side of eqn. (24) is identically zero; implying that $\lambda^* = \lambda^0(Z^*)$.

Applying Taylor's expansion to the RHS of eqn. (36), about the point Z^*, gives the following linearized equation

$$\frac{dZ}{dt} = \varphi(Z) + \varphi_Z(Z)y. \tag{38}$$

Comparing eqn. (38) with the RHS of eqn. (23) yields

$$\frac{dZ}{dt} = \varphi(Z^*) + \varphi_Z(Z^*)(Z - Z^*). \tag{39}$$

Recall that

$$L(Z, \lambda) = J(Z) + \lambda^T g(Z), \tag{40}$$
$$L_Z(Z, \lambda) = J_Z(Z) + \lambda^T g_Z(Z), \tag{41}$$
$$L_{ZZ}(Z, \lambda) = J_{ZZ}(Z) + g_{ZZ}^T(Z)\lambda, \tag{42}$$

then from eqn. (38),

$$\frac{dZ}{dt} = \varphi(Z^*) + \varphi_Z(Z^*)y, \quad [\text{where } y = Z - Z^*]. \tag{43}$$

Suppose $\dot{\varphi}(Z^*) = L_Z(Z, \lambda)$ which is equal to eqn. (41); then differentiating eqn. (41) arrives at

$$\varphi_Z(Z^*) = [J_{ZZ}(Z^*) + \lambda^T g_{ZZ}(Z^*) + g_Z(Z^*)\lambda_Z]y,$$
$$= [L_{ZZ}(Z^*, \lambda^*)) + g_Z(Z^*)\lambda_Z]y. \tag{44}$$

Therefore, eqn. (39) becomes

$$\frac{dZ}{dt} = -[L_Z(Z^*, \lambda^*) + [L_{ZZ}(Z^*, \lambda^*) + g_Z^T(Z^*)\lambda_Z^0(Z^*)](Z - Z^*)], \tag{45}$$

where $\varphi(Z^*) = L_Z(Z^*, \lambda^*)$ and $\varphi_Z(Z^*) = L_{ZZ}(Z^*, \lambda^*) + g_Z^T(Z^*)\lambda_Z^0(Z^*)$.
Considering the unknown coefficient $\lambda_Z^0(Z^*)$ and setting $L_Z(Z^*, \lambda^*) = 0$ yields,

$$\frac{dZ}{dt} = -[L_{ZZ}(Z^*, \lambda^*) + g_Z^T(Z^*)\lambda_Z^0(Z^*)](Z - Z^*)]. \tag{46}$$

We seek to differentiate both sides of eqn. (32) with respect to Z and evaluating at Z^*. Firstly, we differentiate the LHS of eqn. (32) as follows:

$$g_Z(Z^*)\frac{d}{dZ}[J_Z(Z) + \lambda^0(Z)g_Z^T(Z)]$$
$$+ [J_Z(Z) + \lambda^0(Z)g_Z^T(Z)]\frac{d}{dZ}g_Z(Z)$$
$$= g_Z(Z^*)[J_{ZZ}(Z^*) + \lambda_0(Z^*)\frac{d}{dZ}g_Z^T(Z^*) \tag{47}$$
$$+ g_Z^T(Z)\frac{d}{dZ}\lambda^0(Z)]$$
$$+ [J_Z(Z^*)g_{ZZ}(Z^*)\lambda^0(Z)g_Z^T(Z^*)g_{ZZ}(Z^*)$$

and arrives at

$$[J_{ZZ}(Z^*)g_Z(Z^*) + \lambda^0 g_{ZZ}^T(Z^*)]g_Z(Z^*)$$
$$+[J_Z(Z^*)$$
$$+ \lambda_Z^0(Z^*)g_Z^T(Z^*)]g_{ZZ}(Z^*)$$
$$+ \lambda_Z^0(Z^*)g_Z^T(Z^*)g_Z(Z^*). \tag{48}$$

Secondly, we differentiate the RHS of eqn. (32) as follows;

$$\tau\frac{d}{dZ}\left[g_Z^T(Z)g_Z(Z)g(Z)\right]$$
$$= \tau\left[g_Z^T(Z)\frac{d}{dZ}g_Z(Z)g(Z)\right.$$
$$\left.+g_Z(Z)g(Z)\frac{d}{dZ}g_Z^T(Z)\right],$$
$$= \tau[g_Z^T(Z^*)g_Z(Z^*)]g_Z(Z^*)$$
$$+\tau[g_Z^T(Z^*)g_{ZZ}(Z^*)]g(Z^*)$$
$$+ \tau[g_{ZZ}^T(Z^*)]g_Z(Z^*)g(Z^*). \tag{49}$$

Equating eqns. (48) and (49), we have

$$[J_{ZZ}(Z^*)g_Z(Z^*) + \lambda^0 g_{ZZ}^T(Z^*)]g_Z(Z^*) + [J_Z(Z^*)$$
$$+ \lambda_Z^0(Z^*)g_Z^T(Z^*)]g_{ZZ}(Z^*) + \lambda_Z^0(Z^*)g_Z^T(Z^*)g_Z(Z^*) \tag{50}$$
$$= \tau[g_Z^T(Z^*)g_Z(Z^*)]g_Z(Z^*)$$
$$+ \tau[g_Z^T(Z^*)g_{ZZ}(Z^*)]g(Z^*)$$
$$+ \tau[g_{ZZ}^T(Z^*)]g_Z(Z^*)g(Z^*). \tag{51}$$

Applying eqns. (18) and (19) on eqn. (51), we have

$$g_Z(Z^*)[L_{ZZ}(Z^*, \lambda^*) + g_Z^T(Z^*)\lambda_Z^0(Z^*)]$$
$$= \tau g_Z(Z^*)g_Z^T(Z^*)g_Z(Z^*), \tag{52}$$

where $L_{ZZ}(Z^*, \lambda^*) = J_{ZZ}(Z^*)g_Z(Z^*) + \lambda^0 g_{ZZ}^T(Z^*)$ and $L_Z(Z^*, \lambda) = J_Z(Z^*) + \lambda_Z^0(Z^*)g_Z^T(Z^*))$.
Eqn. (52) becomes

$$g_Z(Z^*)L_{ZZ}(Z^*, \lambda^*) + G(Z^*)\lambda_Z^0(Z^*)$$
$$= \tau G(Z^*)g_Z(Z^*), \tag{53}$$

where $G(Z^*) = g_Z^T(Z^*)g_Z(Z^*)$.
Making $\lambda_Z^0(Z^*)$ the subject of eqn. (53) yields

$$\lambda_Z^0(Z^*) = -G^{-1}(Z^*)g_Z(Z^*)L_{ZZ}(Z^*, \lambda^*) + \tau g_Z(Z^*). \tag{54}$$

Recall that

$$y = Z - Z^*, \tag{55}$$
$$\Rightarrow \quad \frac{dy}{dt} = \frac{dZ}{dt} - \frac{dZ^*}{dt}, \tag{56}$$
$$\text{and} \quad \frac{dy}{dt} = \frac{dZ}{dt}. \tag{57}$$

Substituting eqn. (54) into eqn. (46) and considering eqn. (57) yields

$$\frac{dy}{dt} = -[L_{ZZ}(Z^*, \lambda^*) +$$
$$g_Z^T(Z^*)(-G^{-1}(Z^*)g_Z(Z^*)L_{ZZ}(Z^*, \lambda^*)$$
$$+\tau g_Z(Z^*))]y, \tag{58}$$

$$\frac{dy}{dt} = -[L_{ZZ}(Z^*, \lambda^*) +$$
$$+g_Z^T(Z^*)g_z(Z^*)L_{ZZ}(Z^*, \lambda^*)G^{-1}(Z^*)$$
$$-\tau g_Z(Z^*)g_Z^T(Z^*)]y, \tag{59}$$

$$\frac{dy}{dt} = [-L_{ZZ}(Z^*, \lambda^*)[I - g_Z^T(Z)G^{-1}(Z)g_Z(Z)]$$
$$-\tau g_Z^T(Z)g_Z(Z)]y. \tag{60}$$

Since Z^* is zero at equilibrium point, then

$$\frac{dy}{dt} = -[(I - g_Z^T(Z)G^{-1}(Z)g_Z(Z))L_{ZZ}(Z^*, \lambda^*)$$
$$+ \tau g_Z^T(Z)g_Z(Z)]y. \tag{61}$$

Let

$$S(Z) = I - g_Z^T(Z)G^{-1}(Z)g_Z(Z), \tag{62}$$
$$V(Z) = g_Z(Z)g_Z^T(Z), \tag{63}$$

where I is the $N \times N$ identity matrix. Therefore, eqn. (61) becomes

$$\frac{dy}{dt} = -[S(Z^*)L_{ZZ}(Z^*, \lambda^*) + \tau V(Z^*)]y. \tag{64}$$

From above, eqns. (62) and (63) are orthogonal where eqn. (62) is an operator which projects a vector to the null space of $g_Z(Z)$, because $g_Z(Z)S(Z) = 0$.
Multiplying eqn. (64) from the left by $g_Z(Z^*)$ yields

$$g_Z(Z^*)\frac{dy}{dt} = -[\tau g_Z(Z^*)V(Z^*)]y. \tag{65}$$

Eqn. (65) is equivalent to

$$\frac{dy}{dt} = -[\tau V(Z^*)]y. \tag{66}$$

Hence, since $\tau > 0$, it remains to show that $V(Z^*)$ is positive definite. Recall that $V(Z^*) = g_Z^T(Z^*)g_Z(Z^*)$ and $g_Z(Z^*) = G$; where G whose dimension is $N \times (2N + 1)$ is given by eqn. (3.3). Therefore, $Q = GG^T$ with dimension $N \times N$ is given by

$$Q = \begin{bmatrix} 1 + d^2 & -c & 0 & \cdots & 0 \\ -c & \alpha & -c & 0 & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -c \\ 0 & \cdots & 0 & -c & \alpha \end{bmatrix}, \tag{67}$$

where $\alpha = 1 + c^2 + d^2$.

Theorem 3.4 (Sylvester's Criterion). *A real, symmetric matrix is positive definite if and only if all the principal minors are positive definite. Gilbert [7].*

Considering the implementation of **Theorem 3.4** on the matrix Q, we seek to establish that the matrix is positive definite. Firstly, Q is real since for every $q_{i,j} \in Q$, $q_{i,j} \in R$ (i.e. all the diagonal entries in eqn. (67) are real).
Secondly, $Q = g_Z^T(Z^*)g_Z(Z^*)$ is symmetric since $(Q)^T = [g_Z^T(Z^*)g_Z(Z^*)]^T$ $g_Z^T(Z^*)g_Z(Z^*) = Q$.
However, we also seek to establish that the principal minors Q_i, $\forall i = 1, 2, 3, ...N$, of Q are all positive definite. When $i = 1, 2, 3$, and 4, Q_1, Q_2, Q_3 and Q_4 are 1, $1 + c^2$, $1 + c^2 + c^4$ and 1 respectively; and are all positive $\forall c, d > 0$. We then assumed, by mathematical induction, that it is true for Q_k when $i = k$ and Q_{k+1} when $i = k + 1$. And by Cholesky, it was shown that Q_{k+1} is positive definite and that all the principal minors of Q are positive definite [*See Appendix*]. We, therefore, concluded that the matrix $Q = GG^T$ is positive definite.

3.5. Construction of the Modified Gradient Flow Method (MGFM)

We now considered the discretization of the formulation of eqns. (23) and (24) using three level implicit time discretization scheme with splitting parameters θ_1 and θ_2 where $\theta_3 = (1 - \theta_1 - \theta_2)$.
Let $0 = t_0 < t_1 < t_2 < t_3 < ... < t_k = t$ be a sequence time points for the time $t \geq t_0$, $h_k = t_{k+1} - t_k$ and $\delta_{Z_{k+1}} = Z_{k+2} - Z_k$ being the sequence of solution distance between two successive

equilibrium points. We discretize eqn. (23) by the following implicit time stepping algorithm;

$$\frac{Z_{k+2} - Z_{k+1}}{h_{k+1}} = -[\theta_3 L_Z(Z_{k+2}, \lambda_{k+2}^0) + \theta_2 L_Z(Z_{k+1}, \lambda_{k+1}^0) + \theta_1 L_Z(Z_k, \lambda_k^0)]. \tag{68}$$

Simplified to

$$Z_{k+2} - Z_{k+1} = -h_{k+1}[\theta_3 L_Z(Z_{k+2}, \lambda_{k+2}^0) + \theta_2 L_Z(Z_{k+1}, \lambda_{k+1}^0) + \theta_1 L_Z(Z_k, \lambda_k^0)]. \tag{69}$$

Employing Taylor's expansion about $(Z_{k+2}, \lambda_{k+2}^0)$ arrives at

$$Z_{k+1} = Z_k + h_k Z_k + O\|h_k\|^2, \tag{70}$$

$$L_Z(Z_{k+2}, \lambda_{k+2}^0) = L_Z(Z_k, \lambda_k^0) + H(Z_k, \lambda_k^0)\delta_{Z_{k+1}} + O\|\delta_{Z_{k+1}}\|^3, \tag{71}$$

$$L_Z(Z_{k+1}, \lambda_{k+1}^0) = L_Z(Z_k, \lambda_k^0) + H(Z_k, \lambda_k^0)\delta_{Z_k} + O\|\delta_{Z_k}\|^3, \tag{72}$$

where, $\delta_{Z_k} = Z_{k+1} - Z_k$ and
$H(Z_k, \lambda_k) = L_{ZZ}(Z_k, \lambda_k^0) + L_{Z\lambda_k^0}((Z_k, \lambda_k))\lambda_Z(Z_k)$
is the Hessian of $L(Z, \lambda)$. Substituting eqns. (70), (71) and (72) into eqn. (69) yields

$$\begin{aligned} Z_{k+2} = {} & Z_k + h_k Z_k' + O\|h_k\|^2 \\ & - h_{k+1}[\theta_3(L_Z(Z_k, \lambda_k^0) \\ & + H(Z_k, \lambda_k^0)\delta_{Z_{k+1}} + O\|\delta_{Z_{k+1}}\|^3) \\ & + \theta_2(L_Z(Z_k, \lambda_k^0) + H(Z_k, \lambda_k^0)\delta_{Z_k} \\ & + O\|\delta_{Z_k}\|^2) + \theta_1 L_Z(Z_k, \lambda_k^0)]. \end{aligned} \tag{73}$$

Neglecting the higher order terms in eqn. (73) and solving the resultant equation for Z_{k+2} yields

$$\begin{aligned} Z_{k+2} = {} & Z_k + h_k Z_k - h_{k+1}\theta_3 L_{Zk} \\ & - h_{k+1}\theta_3 H_{Zk}Z_{k+2} + h_{k+1}\theta_3 H_{Zk}Z_k \\ & - h_{k+1}\theta_2 L_{Zk} + + h_{k+1}h_k\theta_2 H_{Zk}L_{Z_k} \\ & - h_{k+1}\theta_1 L_{Z_k}. \end{aligned} \tag{74}$$

For equal step size, $h_{k+1} = h_k$ with $\theta_1 = 0$ we obtain Z_{k+2} expressed below

$$\begin{aligned} Z_{k+2} = {} & Z_k + h_k Z_k - h_k\theta_3 L_{Zk} - h_k\theta_3 H_{Zk}Z_{k+2} \\ & + h_k\theta_3 H_{Zk}Z_k - h_k\theta_2 L_{Zk} \\ & + h_k^2\theta_2 H_{Zk}L_{Zk}. \end{aligned} \tag{75}$$

In collecting like terms and factorizing eqn. (75), it becomes

$$\begin{aligned} (I + {} & h_k\theta_3 H_{Zk})Z_{k+2} \\ & = (I + h_k\theta_3 H_{Z_k})Z_k - (\theta_2 + \theta_3)h_k L_{Zk} \\ & + h_k^2\theta_2 H_{Zk}L_{Zk}, \\ & = (I + h_k\theta_3 H_{Z_k})Z_k + (-h_k I \\ & + h_k^2\theta_2 H_{Zk})L_{Zk}. \end{aligned} \tag{76}$$

Making Z_{k+2} the subject of formula in eqn. (76) yields

$$\begin{aligned} Z_{k+2} = {} & (I + h_k\theta_3 H_{Zk})^{-1}[(I + h_k\theta_3 H_{Zk})Z_k \\ & + (-h_k I + h_k^2\theta_2 H_{Zk})L_{Zk}], \\ = {} & Z_k - h_k(I + h_k\theta_3 H_{Zk})^{-1} \\ & \times [I - h_k\theta_2 H_{Zk}]L_{Z_k}, \end{aligned}$$

for any $\theta_2, \theta_3 \in [0, 1]$ and $\theta_2 + \theta_3 = 1$ by the time-stepping algorithm for convex functions.
Also, from eqn. (76), we obtained

$$\begin{aligned} (I + {} & h_k\theta_3 H_{Z_k})(Z_{k+2} - Z_k) \\ & = -h_k(I + h_k\theta_2 H_{Zk})L_{Zk}. \end{aligned} \tag{77}$$

Given that $\delta_{k+1} = Z_{k+2} - Z_k$ and $\theta_2 = 0$, eqn. (77) becomes

$$(I + h_k\theta_3 H_{Z_k})\delta_{Z_{k+1}} = -h_k L_{Zk}. \tag{78}$$

Thus, for any initial guess Z_0, eqn. (77) defines a series converging to the solution of the optimal control problem stated in eqns. (1) to (3). Hence, the Modified Gradient Flow (MGF) Algorithm for the Discretized Optimal Control Problem (OCP) is stated as follows:

MGF Algorithm for Discretized OCP

Step 1: **Input** the operator $G \in R^{N\times(2N+1)}$,
$Z_0 \in R^{2N+1}$, parameters $\theta_2, \theta_3 \in [0, 1]$,
sequence of time step-size $\{h_k\}$ and tolerance
$(\sigma > 0)$ sufficiently small. Set $k = 0$.
Step 2: **Compute** $\lambda_{Z_k}^0$ using eqn. (54).
Step 3: **Solve** for $\delta_{Z_{k+1}}$ in the system using
the CGM $(I + h_k\theta_3 H_{Z_k})\delta_{Z_{k+1}} = -h_k L_{Z_k}$
(where $\delta_{Z_{k+1}} = Z_{k+2} - Z_k$)
Step 4: **Update**: $Z_{k+2} = Z_k + \delta_{Z_{k+1}}$.
Step 5: **Stop** the process if $\|H_{Z_k}\| \le \sigma$,
otherwise, set $k = k + 1$ and repeat step 3.

3.6. Error and Convergence Analysis

The convergence analyses of the MGFM requires us to shown the rate and type of convergence of the algorithm for varying values of parameters $\theta 1, \theta 2$ and $\theta 3$. We shall show that the algorithm is super-linearly convergent for some θ_2 and θ_3 and quadratically convergent if $\theta_1 = \theta_2 = 0$ and $\theta_3 = 1$.

Theorem 3.5. *Let $\{Z_k\}$ be the sequence defined in eqn. (77) and $(Z^{\aa}, \lambda^{\aa})$ be the solution of eqns. (1) and (2). If the initial guess Z_0 is sufficiently close to Z, then we have*

1. *Z_k converges linearly to $Z^{\aa}$ if $\theta_1, \theta_2, \theta_3 \in [0, 1]$ and $h_k > 0$ is sufficiently small;*

2. *Z_k converges to $Z^{\aa}$ quadratically if $\theta_2 = 0$, $\theta_3 = 1$ and $h_k \to \infty$.*

We shall prove the two cases separately.

- **Case 1.** $\theta_1, \theta_2, \theta_3 \in [0, 1]$.

  From eqn. (77) we have: $(I + h_k\theta_3 H_{Z_k})(Z_{k+2} - Z_k) = -h_k(I - h_k\theta_2 H_{Z_k})L_{Z_k}$.

  Hence,

  $$Z_{k+2} = Z_k - h_k[(I - h_k\theta_2 H_{Z_k})L_{Z_k} + \theta_3 H_{Z_k}(Z_{k+2} - Z_k)].$$

  Let (Z^*, λ^*) be the minimum points of eqns. (16) and (17) and $e_{k+1} = Z_{k+1} - Z^*$.

  Subtracting Z^* from both sides of the above equation and noting that we have

  $$
  \begin{aligned}
  e_{k+1} &= e_k - h_k[(I - h_k\theta_2 H_{Z_k})(L_{Z_k} - L_{Z_k^*}) \\
  &\quad + \theta_3 H_{Z_k}(e_{k+1} - e_k)].
  \end{aligned}
  \tag{79}
  $$

  Thus, by the Mean Value Theorem,

  $$
  \begin{aligned}
  e_{k+1} &= e_k - h_k[(I - h_k\theta_2 H_{Z_k})H_{Z_k}^{\zeta_k} e_k \\
  &\quad + \theta_3 H_{Z_k}(e_{k+1} - e_k)],
  \end{aligned}
  \tag{80}
  $$

  where ζ_k denotes a point in the line segment connecting Z_k and Z^*. Taking the norm of both sides, we have

  $$\|e_{k+1}\| \leq \gamma(Z_k, \lambda_k^0, \zeta_k, \theta_1, \theta_2, \theta_3, h_k)\|e_k\|, \tag{81}$$

  where

  $$
  \begin{aligned}
  &\gamma(Z_k, \lambda_k^0, \zeta_k, \theta_1, \theta_2, \theta_3, h_k) = \\
  &\|\{I - h_k(I + \theta_3 H_{Z_k})^{-1}(I - \theta_2 H_{Z_k})H_{Z_k}^{\zeta_k}\}\|.
  \end{aligned}
  \tag{82}
  $$

  From eqn. (81), we see that if $\gamma(Z_k, \lambda_k^0, \zeta_k, \theta_1, \theta_2, \theta_3, h_k) < 1$, then e_k converges to zero linearly. To show this, we need to prove that H_{Z_k} is positive definite. But by theorems and , H_{Z_k} is positive definite. With these, when Z_k is sufficiently close to Z^*, from eqn. (82), it is seen that if $\theta_1, \theta_2, \theta_3 \in [0, 1]$ and h_k sufficiently small, we have

  $$
  \begin{aligned}
  &\gamma(Z_k, \lambda_k^0, \zeta_k, \theta_1, \theta_2, \theta_3, h_k) \\
  &\leq \left(1 - \frac{h_k(1 - h_k\theta_2\lambda_{min}^k)\lambda_{min}^k}{1 + h_k\theta_3\lambda_{max}^k}\right) < 1,
  \end{aligned}
  \tag{83}
  $$

  where λ_{min}^k and λ_{max}^k denote the minimum and maximum eigenvalues of H_{Z_k}, respectively. Therefore, eqn. (81) implies that $\lim_{k\to\infty} e_k = 0$ is linear. This proves that Z_k converges linearly to $Z^{\mathring{a}}$ if $\theta_1, \theta_2, \theta_3 \in [0, 1]$ and $h_k > 0$ is sufficiently small.

- **Case 2.** $\theta_3 = 1$.

  From eqn. (77), when $\theta_3 = 1$ and $h_k = h$, we have

  $$\frac{Z_{k+2} - Z_k}{h} = -(L_{Z_k} + H_{Z_k}\delta_{Z_{k+1}}).$$

  When $h \to \infty$, the above equation becomes

  $$H_{Z_k}\delta_{Z_{k+1}} + L_{Z_k} = 0. \tag{84}$$

  Eqn. (84) coincides with the equation resulting from the application of the Newton's method to $L_Z = 0$. Hence, Z_k converges quadratically to Z^* if Z_0 is sufficiently close to Z^*.
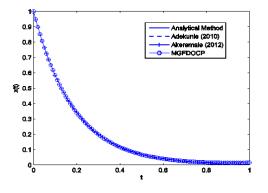

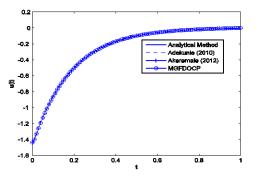Figure 1. State Behaviors in various Algorithms

Figure 2. Control Behaviors in various Algorithms

4. Examples and Numerical Results

The numerical examples and respective results are presented below based on the construction of modified gradient flow method (MGFM) for discretized optimal control problems. The results show the robustness and efficiency of the scheme (MGFM) as compared with existing schemes.

Example 1

Considering the optimal control problem presented by [1]

$$\text{Minimize} \quad I(x, u) = \int_0^1 (x^2(t) + u^2(t))dt, \tag{85}$$

$$\text{subject to} \quad \dot{x} = 2x(t) + 5u(t), \quad x(0) = 1. \tag{86}$$

The analytical solution is given as

$$\begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} 1.0000 \\ -0.5908 \end{pmatrix} e^{-5.3852t}, \tag{87}$$

and the control variable is given as $u(t) = -1.4770e^{-5.3852t}$. The solution of eqns. (85) and (86) were obtained, using $h = 0.1$ and $h_k = 20$, by the modified gradient algorithm for discretized optimal control problems and the results were compared with the existing algorithms as shown in table 1 below. However, the step-size $h = 0.1$ was chosen for the purpose of illustration. Meanwhile smaller choices of step-size could be used for better refinement.

Algorithm	objective functional value
Analytical Method	0.2953894
Adekunle [1]	0.3033885
Akeremale [2]	0.2955932
MGFM	0.2955879

Table 1: Comparison of Objective Functional Values of Existing Methods for Example 1

Results: From **table 1** above, it was observed that the objective functional value (0.2955879) for the developed MGFM agreed favourably with the value (0.2953894) obtained from the analytical method and better than the values obtained by Adekunle [1] and Akeremale [2] from Conjugate Gradient Method (CGM) and Extended Conjugate Gradient Method (ECGM) respectively.

Table 2a: Errors in the State variables for Example 1

t	$\|x(t) - \hat{x}(t)\|$ $(\times 10^{-3})$		
	Adekunle [1]	Akeremale [2]	MGFM
0.0	0	0	0
0.1	3.2204	0.5528	0.5525
0.2	3.8300	0.3680	0.3677
0.3	3.4841	0.3414	0.3410
0.4	2.9424	0.4447	0.4442
0.5	2.5506	0.6968	0.6962
0.6	2.4892	1.1667	1.1650
0.7	2.9118	1.9867	1.9858
0.8	4.0417	3.4014	3.4005
0.9	6.2660	5.8292	5.8279
1.0	10.266	9.9981	9.9789

t	$\|u(t) - \hat{u}(t)\|$ $(\times 10^{-4})$		
	Adekunle [1]	Akeremale [2]	MGFM
0.0	38.796	38.419	38.4202
0.1	18.131	0.6668	0.6555
0.2	7.9765	0.2654	0.2562
0.3	3.1973	0.0149	0.0755
0.4	1.1303	0.1891	0.1949
0.5	4.3839	0.4248	0.4294
0.6	4.8076	0.7726	0.7761
0.7	1.0193	1.3417	1.3442
0.8	2.0619	2.3045	2.3062
0.9	3.8245	3.9497	3.9505
1.0	6.7705	6.4026	6.4027

Table 2 above shows the errors in the state and control variables for **example 1** while the errors generated with the MGFM were minimal compared with the errors generated from Adekunle [1] and Akeremale [2]. **Figures 1** and **2** below clearly show the

comparisons of the behaviors of the state and control trajectories to the various algorithms respectively. It was observed that the trajectories for both the state and control variables overlap due to closeness in results.

Example 2

Considering the optimal control problem presented by [1]

$$\text{Minimize} J(x, u) = \int_0^1 (x^2(t) + u^2(t))dt, \qquad (88)$$

$$\text{subject to} \quad \dot{x} = 1.705x(t) + 3.021u(t), \qquad (89)$$

$$x(0) = 1. \qquad (90)$$

Table 3: Comparison of Objective Functional Values of Existing Methods for Example 2

Table 4a: Errors in the State variables for Example 1

Algorithm	Objective Functional Value
Analytical Method	0.5647
Adekunle [1]	0.5746
Akeremale [2]	0.5649
MGFM	0.5648

t	$\|x(t) - \hat{x}(t)\|$ $(\times 10^{-4})$		
	Adekunle [1]	Akeremale [2]	MGFM
0.0	10.000	10.0000	1.0000E
0.1	3.2204	5.528	5.5253
0.2	2.0509	4.347	3.9394
0.3	3.3242	2.604	2.3946
0.4	3.2376	2.435	2.2688
0.5	2.9852	2.646	2.4997
0.6	2.6832	3.2433	3.0992
0.7	2.3921	4.2860	4.1303
0.8	2.1365	5.8933	5.7138
0.9	1.9144	8.2567	8.0424
1.0	1.7012	10.716	11.403

Table 4b: Errors in the Control variables for Example 1

Results: From **table 3** above, it was observed that the objective functional value, 0.5648, for MGFM agreed favourably with the value (0.5647) obtained from the analytical method; and better than the values obtained by Adekunle [1] and Akeremale [2] from Conjugate Gradient Method (CGM) and Extended Conjugate Gradient Method (ECGM) respectively as expressed in table 3 above.

Results: Table 4 above shows the errors in the state and control variables for **example 2** where the errors generated from the use of MGFM were minimal compared to the errors generated from same examples in Adekunle [1] and Akeremale [2]. **Figures 3** and **4** below clearly show the comparisons of the behaviors of the state and control trajectories to the various algorithms

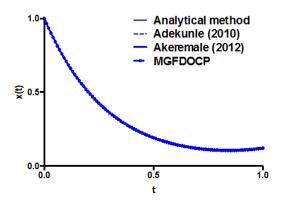| t	$\|u(t) - \hat{u}(t)\| (\times 10^{-4})$		
	Adekunle [1]	Akeremale [2]	MGFM
0.0	0.2928	0.3404	0.2893
0.1	1.8337	0.5220	0.3179
0.2	1.0622	1.8094	2.4019
0.3	5.9872	1.1203	2.3334
0.4	3.3442	0.4000	0.4941
0.5	1.9971	0.7234	0.7965
0.6	1.5137	1.1255	1.1823
0.7	1.6411	1.6579	1.7019
0.8	2.2516	2.3872	2.4210
0.9	3.3101	3.4037	3.4311
1.0	2.9285	3.4045	2.8932

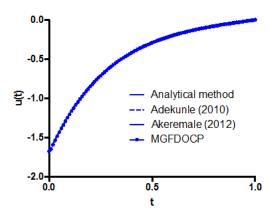Figure 3. State Behaviors in various Algorithms

Figure 4. Control Behaviors in various Algorithms

respectively. It was observed that the trajectories for both the state and control variables overlap due to closeness in results.

5. Conclusion

This paper developed a new scheme for the solution of a quadratic optimal control problem with linear equality constraint. The method is a gradient flow formulation that requires dis-

cretization techniques and a splitting parameter approach for the gradient of the equation whose solution converges to a local minimum of the original problem; either linearly or quadratically. Two linear equality constrained optimal control problems were considered and their results compared with the analytical and two known existing schemes. Thus, it was observed that the new scheme compared favorably with the analytic solutions and performed better than the Conjugate Gradient Method used in [2] in terms of convergence and accuracy. It is therefore obvious that this developed algorithm, due to its robustness and efficiency, would provide a new platform for solving optimal control problems with or without box constraints, equality or inequality constraints or with delay or non-delay dynamical systems.

References

[1] A. I. Adekunle, "Algorithm for a Class of Discretized Optimal Control Problems", M.Tech. Thesis, Federal University of Technology, Akure, Nigeria (2011) (Unpublished).

[2] O. C. Akeremale, "Optimization of Quadratic Constrained Optimal Control Problems using Augmented Lagrangian Method", M.Tech. Thesis, Federal University of Technology, Akure, Nigeria (2012) (Unpublished)

[3] W. Behrman, "An Efficient Gradient Flow Method for Unconstrained Optimization", PhD Thesis, Stanford University (1998) (Unpublished).

[4] J. T. Betts, "Practical Methods for Optimal Control Problem Using Non linear programming", SIAM, Philadelphia, (2001).

[5] Y. Evtushenko, "Generalized Lagrange Multipliers Technique for Nonlinear Programming", JOTA, **21** (1977) 121.

[6] Y. G. Evtushenko & V. G. Zhadan, "Stable Barrier Projection and Barrier Newton Methods", Nonlinear Programming Optimization Methods and Software, **3** (1994) 237.

[7] G. T. Gilbert, "Positive Definite Matrices and Sylvester's Criterion", The American Mathematical Monthly, Taylor & Francis, **98** (1991) 44.

[8] W. M. Haddad, S. G. Nersesov & V.S. Chellaboina, "Lyapunov Function Proof of Poincare's Theorem", International Journal of Systems Science, **35** (2004) 287.

[9] J. B. Layton, "Efficient direct computatioion of the Pseudo-inverse and its gradient", International Journal for Numerical Methods in Engineering, **40** (1997) 4211.

[10] W. H. Morris, S. Stephen & L. D. Robert, "Differential Equations, Dynamical Systems, and an Introduction to Chaos", Elsevier Academic Press, USA (2004) 194.

[11] O. Olotu & K. A. Dawodu, "Quasi-Newton Embedded Augmented Lagrangian Algorithm for Discretized Optimal Proportional Control Problems", Journal of Mathematical Theory and Modeling, **3** (2013a) 67.

[12] O. Olotu & K. A. Dawodu, "On the Discretized Algorithm for Optimal Proportional Control Problems Constrained by Delay Differential Equations", Journal of Mathematical Theory and Modeling, **3** (2013b) 157.

[13] O. Olotu & S. A. Olorunsola, "An Algorithm for a Discretized Constrained, Continuous Quadratic Control Problem", Journal of Applied Sciences, **8** (2006) 6249.

[14] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze & E. F. Mishchenko, "The Mathematical Theory of Optimal Processes", Interscience Publishers, London (1962).

[15] S. Wang, X. Q. Yang K. L. Teo "A Unified Gradient Flow Approach to Constrained Nonlinear Optimization Problems", Computational Optimization and Applications, **25** (2003) 251.

APPENDIX

Let Q_i represent the leading principal minor of Q, $\forall i = 1, 2, 3, ..., N$ of the matrix $Q = GG^T$; we then considered the determinants of the principal minors $Q_1, Q_2, Q_3, ...$ as follows:

For $i = 1$, $Q_1 = 1 + d^2 > 0$, $\forall d$.

For $i = 2$,

$$Q_2 = \begin{bmatrix} 1 + d^2 & -c \\ -c & 1 + c^2 + d2 \end{bmatrix}, \tag{91}$$

such that $det(Q_2) = 1 + 2d^2 + c^2 d^2 + d^4 > 0$ for all c, d.

For $i = 3$,

$$Q_3 = \begin{bmatrix} 1 + d^2 & -c & 0 \\ -c & 1 + c^2 + d^2 & -c \\ 0 & -c & 1 + c^2 + d^2 \end{bmatrix}, \tag{92}$$

such that $det(Q_3) = 1 > 0 \ \forall c$.

Therefore, $|Q_3|$ is positive.

For $i = 4$,

$$Q_4 = \begin{bmatrix} 1 + c^2 & -c & 0 & -d \\ -c & 1 + c^2 & -c & 0 \\ 0 & -c & 1 & 0 \\ -d & 0 & 0 & 2d^2 \end{bmatrix}, \tag{93}$$

such that $det.(Q_4) = d^2 > 0$. $\forall c, d$.

Therefore, $det(Q_4) = d^2 > 0$, $\forall d$, i.e. $|Q_4|$ is positive. Hence, by mathematical induction, since it is true for values of Q_i, $i = 1, 2, 3, 4$, then we assume that it is true for $i = k$, i.e. Q_k, while we shall prove that it is true for $i = k + 1$, i.e. Q_{k+1}. Set $\sigma^2 = 1 + c^2 > 0$, $c^2 > 0$, $1 > 0$, $2d^2 > 0$ and $d^2 > 0$ where $1 + c^2$, c^2, 1, $2d^2$ and d^2 are principal diagonals of Q.

Thus,

$$Q_{k+1} = \begin{bmatrix} Q_k & \beta_{k+1} \\ \beta_{k+1}^T & \sigma^2 \end{bmatrix}, \tag{94}$$

where

$$\beta_{k+1} = \begin{bmatrix} Q_{1,k+1} & Q_{2,k+1} & \dots & Q_{k,k+1} \end{bmatrix}^T. \tag{95}$$

By Cholesky, Q_{k+1} is said to be positive definite if there exists a lower triangular matrix $L_{i,j}$ such that $Q_{k+1} = LL^T$. That is,

$$\begin{bmatrix} Q_k & \beta_{k+1} \\ \beta_{k+1}^T & c^2 \end{bmatrix}$$

$$= \begin{bmatrix} H_k & 0 \\ L_{k+1,1}^T & L_{k+1,k+1} \end{bmatrix} \begin{bmatrix} H_k^T & L_{k+1,1}^T \\ 0 & L_{k+1,k+1} \end{bmatrix},$$

where H_k is a lower triangular matrix with positive diagonal entries. However, eqn. (5) becomes

$$\begin{bmatrix} Q_k & \beta_{k+1} \\ \beta_{k+1}^T & c^2 \end{bmatrix} =$$

$$\begin{bmatrix} H_k H_k^T & H_k L_{k+1,1} \\ L_{k+1,1}^T H_k^T & L_{k+1,1}^T L_{k+1,1} + L_{k+1,k+1}^2 \end{bmatrix}, \tag{96}$$

such that $Q_k = H_k H_k^T$, $\beta_{k+1}^T = L_{k+1,1}^T H_k^T$, $\beta_{k+1} = H_k L_{k+1,1}$ and $\sigma^2 = L_{k+1,1}^T L_{k+1,1} + L_{k+1,k+1}^2$.

Thus, since the diagonal entries of H_k are greater than zero, it is non-singular and therefore the linear system of the equation has a unique solution given by $L_{k+1,1} = H_k^{-1} \beta_{k+1}$ and a positive value for $L_{k+1,k+1}$ can be obtained if $\sigma^2 - L_{k+1}^T L_{k+1} > 0$. Hence,

$$\begin{aligned} 0 < |Q_k| &= det \begin{bmatrix} Q_k & \beta_{k+1} \\ \beta_{k+1}^T & \sigma^2 \end{bmatrix}, \\ &= Q_k[\sigma^2 I - \beta_{k+1}^T (Q_k)^{-1} \beta_{k+1}]. \end{aligned}$$

Substituting the values of β_{k+1} and Q_k, we have

$$\begin{aligned} Q &= Q_k[\sigma^2 I - (H_k L_{k+1,1})^T (H_k H_k^T)^{-1} H_k L_{k+1,1}], \\ &= Q_k \sigma^2 - L_{k+1,1}^T L_{k+1,1}, \end{aligned}$$

where

$$\begin{aligned} (H_k H_k^T)^{-1} &= Q_k^{-1}, \\ Q_k &= [\sigma^2 - L_{k+1,1}^T L_{k+1,1}]. \end{aligned}$$

Since $det Q_k > 0$, it follows that $\sigma^2 - L_{k+1,1}^T L_{k+1,1} > 0$. Hence,

$$L_{k+1,k+1} = \sqrt{\sigma^2 - L_{k+1,1}^T L_{k+1,1}}. \tag{97}$$

Thus $L_i > 0$ arises from the fact that G is positive definite (LICQ). Hence, the pair $(Z(t), \lambda^0(Z(t)))$ tends to (Z^*, λ^*) as $t \to \infty$ provided the starting point Z_0 is sufficiently close to point Z^*.